Pipelines de ETL e Machine Learning com Apache Spark

Bibliografia, Referências e Links Úteis

Core Default

https://hadoop.apache.org/docs/r3.3.6/hadoop-project-dist/hadoop-common/core-default.xml

HDFS Default

https://hadoop.apache.org/docs/r3.3.6/hadoop-project-dist/hadoop-hdfs/hdfs-default.xml

YARN Default

https://hadoop.apache.org/docs/r3.3.6/hadoop-yarn/hadoop-yarn-common/yarn-default.xml

Mapred Default

https://hadoop.apache.org/docs/r3.3.6/hadoop-mapreduce-client/hadoop-mapreduce-client-core/mapred-default.xml

Spark Properties

https://spark.apache.org/docs/latest/configuration